DOCUMENT RESUME

ED 242 050 EA 016 570

AUTHOR	Natriello, Gary
TITLE	Evaluation Frequency, Teacher Influence, and the Internalization of Evaluation Processes: A Review of Six Studies Using the Theory of Evaluation and Authority. Final Report.
INSTITUTION	Oregon Univ., Eugene. Center for Educational Policy and Management.
SPONS AGENCY	National Inst. of Education (ED), Washington, DC.
PUB DATE	Nov 83
NOTE	54p.
PUB TYPE	Information Analyses (070)
EDRS PRICE	MF01/PC03 Plus Postage.
DESCRIPTORS	Accountability; Elementary Secondary Education; *Evaluation; *Evaluation Methods; Evaluative Thinking; Teacher Administrator Relationship; Teacher Attitudes; Teacher Behavior; *Teacher Evaluation; *Teacher Improvement; *Teacher Supervision

ABSTRACT

A conceptual framework is developed to consider two dimensions of evaluation systems likely to lead performers to internalize the evaluation process. The theory of evaluation and authority developed by Dornbusch and Scott provides the basis for a review of six studies in support of two propositions: (1) the more frequently performers are evaluated, the more likely they will be to accept or internalize the evaluation process (up to a point); (2) the more influence performers have over evaluation, the more likely they will be to accept or internalize the evaluation process (up to a point). A brief discussion of the theory of evaluation and authority as it applies to schools is followed by a discourse on the rationale for the two propositions cited above. Then the basic features of the six studies supporting these propositions are described. These include two questionnaire studies, two interview studies, and two comparative studies. The next section examines the frequency of evaluation activities as reported by teachers and the degree of influence teachers have over those activities. This is followed by an examination of the relationships between frequency of evaluation, teacher influence over evaluation, and internalization. These propositions are then discussed in terms of their application to the evaluation of students. The paper concludes by identifying directions for further research and theory construction. (TE)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED242050

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

X This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

Evaluation Frequency, Teacher Influence, and the Internalization of Evaluation Processes: A Review of Six Studies Using the Theory of Evaluation and Authority

by

Gary Natriello

November 1983

Center for Educational Policy and Management
College of Education
University of Oregon
Eugene, Oregon 97403
(503) 686-5173

EA 016 570

The preparation of this final report was made possible through an Institutional Grant awarded by the National Institute of Education to the Center for Educational Policy and Management. The opinions expressed in this report do not necessarily reflect the positions or policies of the NIE or the U.S. Department of Education.

TABLE OF CONTENTS

Introduction.....1

1. The Theory of Evaluation and Authority.....4

2. Basic Propositions.....6

   Internalization of the Evaluation Process by Performers.....6

   Dimensions of Evaluation Systems.....6

   Proposition 1: Frequency of Evaluation and Internalization.....7

   Proposition 2: Performer Influence and Internalization.....13

3. Basic Features of the Six Studies.....17

4. Frequency of Evaluation, Teacher Influence Over Evaluations and Indicators of Internalization in Schools.....19

   Frequency of Evaluation.....19

   Teacher Influence Over the Evaluation Process.....25

   Indicators of Teacher Internalization or Acceptance of the Evaluation System.....28

5. The Impact of Frequency of Evaluation and Teacher Influence Over Evaluation Processes on Teacher Acceptance of the Evaluation System.....32

6. A Brief Look at the Impact of Evaluation Processes on Students.35

7. Directions for Further Research and Theory Construction.....37

References.....41

Evaluation Frequency, Teacher Influence, and the Internalization of Evaluation Processe:

A Review of Six Studies Using the Theory of Evaluation and Authority

Introduction

Writing of evaluation as a mechanism to link the work of administrators to the work of teachers, Duckworth (1981) argues that teachers must ultimately internalize the evaluation process by articulating their tasks, criteria, outcome samples, and appraisals. He continues by suggesting that teachers will do this only if it poses no threat to their personal status and if the results not only improve their efficacy but also lead to necessary changes in administrative practices and policies.

This paper will develop a conceptual framework to consider two dimensions of evaluation systems likely to lead performers to internalize the evaluation process. The paper will rely on the an evolving body of theory and research on evaluation as a control process in organizations. The theory of evaluation and authority developed by Dornbusch and Scott (1975) and extended by Natriello and Dornbusch (in press) has provided the basis for a series of studies of the evaluation of teachers and students in schools. Six of these studies are reviewed here in support of two general propositions. The first proposition asserts that there is a curvilinear relationship between the frequency with which performers are evaluated and the extent to which they internalize the evaluation process. The second proposition asserts a curvilinear relationship between the degree of influence performers have over the evaluation process and the extent to which they internalize the process.

The paper proceeds by: 1) briefly discussing the theory of evaluation

and authority as it applies to schools; 2) developing the rationale for the two propositions; 3) describing the basic features of the six studies; 4) examining the frequency of evaluation activities as reported by teachers and the degree of influence teachers report having over those evaluation activities; 5) examining the relationships between frequency of evaluation and teacher influence over evaluation and the indicators of internalization; 6) discussing the propositions as they might apply to the evaluation of students; and 7) identifying directions for future research and theory construction.

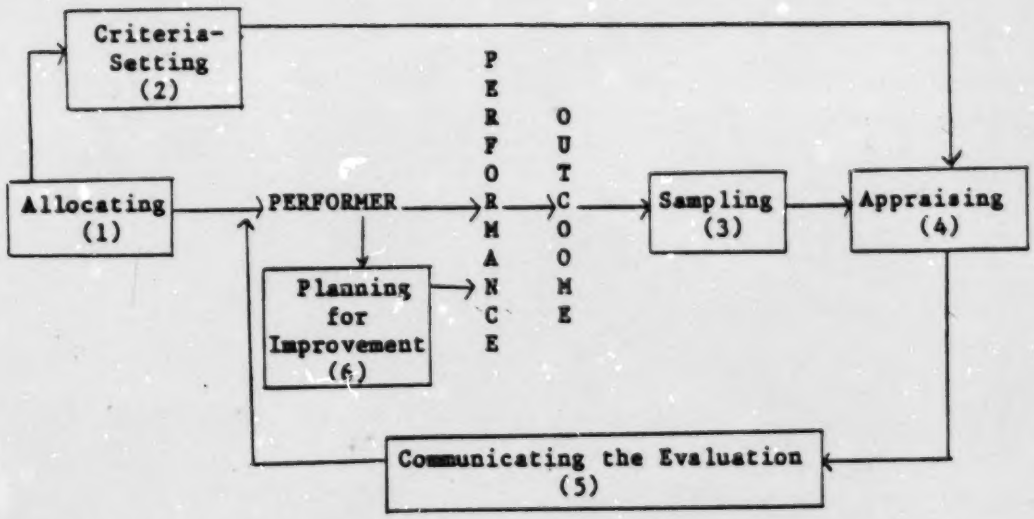
1. The Theory of Evaluation and Authority

The theory of evaluation and authority (Dornbusch and Scott, 1975) and subsequent extensions (Natriello and Dornbusch, 1981; in press) specify stages of evaluation activity in a model of the evaluation process. The form of the model bears a strong resemblance to a servo-mechanism arrangement in which the behavior of a system component is monitored and regulated by means of a feedback loop. The general form of the model makes it particularly useful for orienting a range of studies of evaluation processes in schools, both studies of practices used by administrators for the evaluation of teachers and studies of practices used by teachers for the evaluation of students. This review focuses on the evaluation of teachers and only briefly considers parallel issues in the evaluation of students.

The stages of the model are depicted in Figure 1.

Figure 1

A Model of the Evaluation Process



The six stages of the evaluation model represent a single cycle of the evaluation process. The process begins when a task is allocated or assigned to an individual performer. Determining which properties of the task will be taken into account in evaluating task performance or setting the criteria for the task represents the second stage. The third stage is the process of gathering a sample of information on the actual performance and outcomes of the task. Comparing the information collected on the performance with the evaluative criteria established for the task and assigning an evaluation to the performance are the activities involved in the appraisal process during the fourth stage. Communicating the results of the evaluation of the performance to the individual performer constitutes the fifth stage of the evaluation process. Finally, determining the steps to be taken to insure future performance at an acceptable level (i.e., planning for improvement) is represented by the sixth stage of the model.

The six stages of the model are distinguished because in many organizations they are assigned to different individuals responsible for some part of the evaluation process. For example, in some organizations sampling is handled by designated inspectors; in others the improvement activities are accomplished by referral to employee assistance specialists. Each of these six stages may be evident to a greater or lesser degree in schools. At times these activities are not performed explicitly. At other times these activities are not performed at all. Many of the tasks expected of performers in organizations are never included as part of the formal system for performance evaluation. Even considering both formal and informal evaluations, evaluation may be an infrequent experience for teachers. The model, however, provides a formal analysis of the procedures that would be involved in a performance evaluation system and guides

inquiry into the impact of evaluation processes in schools.

2. Basic Propositions

Internalization of the Evaluation Process by Performers

A basic task of this paper is to identify dimensions of evaluation processes that have an impact on what Duckworth (1981) has referred to as internalizing the evaluation process. For the purposes of this paper "internalization" will be treated as a general rather than a technical term, that is, as a term suggestive of a variety of teacher reactions to the evaluation process. For example, studies using the theory of evaluation and authority have examined the impact of evaluation on teacher satisfaction, teacher perceptions that evaluations accurately reflect their performance, teacher perceptions that the evaluations they receive are helpful, and teacher perceptions of their effort and effectiveness. While each of these variables is in some ways unique, each will be treated as an indicator of the general concept of teacher acceptance or internalization of the evaluation process.

Dimensions of Evaluation Systems

If internalization or acceptance of the evaluation process is a desired outcome, then it is important to understand the dimensions or features of evaluation systems likely to foster such an outcome. Any number of dimensions of evaluation systems have been specified. For example, Meyer, Kay, and French (1965) examined the frequency with which supervisors took samples of the performances and outcomes of subordinates, the amount of criticism conveyed during feedback sessions, and subordinate participation in the process of planning for improvement following less than perfect evaluations. Locke (1975) studied the degree of difficulty and the specificity of goals set for performers. Tannenbaum (1968) and

Landy, Barnes, and Murphy (1978) examined the degree to which supervisors and subordinates shared the same perceptions of tasks and the same values. Szilagyi (1980) focused on the use of punitive sanctions as a result of poor evaluations.

Studies based on the theory of evaluation and authority have also looked at a variety of dimensions of evaluation systems. Dornbusch and Scott (1975) discussed the relationship between evaluations and sanctions, and the importance of the sanctions associated with evaluations to performers. Natriello and Dornbusch (in press) examined the frequency of evaluations, the degree to which standards were challenging, and the extent to which the evaluations received by one performer were consistent with those received by other performers. Clearly, evaluation systems are complex operations with a variety of dimensions suitable for investigation.

In light of the interest in how evaluation might have an impact on the degree to which performers internalize or accept the system, the present analysis will concentrate on two dimensions of evaluation systems: the frequency of evaluation activities experienced by the performer, and the extent to which performers are able to exercise influence over evaluation processes. Both of these dimensions are likely to affect the internalization or acceptance of the evaluation process by performers.

Proposition 1: Frequency of Evaluation and Internalization

There are several good reasons to concentrate on the frequency of evaluation as a dimension of evaluation systems. Not only is frequency of activity a very tangible and easily examined feature of a system, but it is also a good indicator of the supervisory resources that must be committed to the evaluation process. To operate most efficiently an organization should have evaluation activities at intervals that facilitate optimum

performance without incurring the waste of resources, both supervisors' and performers' time, that comes from overly frequent evaluations. Ideally, managers might even wish to be in a position to make decisions about the frequency of evaluation with some knowledge of the additional returns likely to result from increasing that frequency and to weigh those anticipated additional returns against those to be realized from using their time in other ways. Since each organization and each manager will operate in a different context of alternative uses of time, it will never be possible to provide a single recommendation regarding the appropriate frequency of evaluation activity. Nevertheless, considering the impact of frequency of evaluation on performer internalization and acceptance will lead to the development of general guidelines for supervisors.

The activities associated with the evaluation of performance in organizations have two types of effects which might be referred to as rational or operating effects and social or symbolic effects. Each of these types of effects suggests the same relationship between frequency of evaluation and internalization of the process by performers. However, since each type of effect provides a unique perspective for considering the appropriate frequency of evaluations, each will be reviewed separately.

Rational effects refer to those effects that arise from the internal operation of the system. Performers will expect the evaluation system to operate in ways that make rational sense. For example, performers will expect that the evaluations they receive in the feedback stage of the evaluation process have some relationship to their performance of the assigned tasks. When this is not the case, performers will find it difficult to internalize and accept the system. The frequency with which evaluation activities are performed will have a definite impact on the ability of performers to perceive the system as rational.

If evaluation activity is very infrequent, performers are less likely to perceive the connection between their past performance and their evaluations. Activities performed six months in the past are likely to appear irrelevant and perhaps even arbitrary when used as the basis for evaluations received today. Thus, very infrequent evaluation activity is less likely to be internalized and accepted by performers.

But what of very frequent evaluation activity? If very infrequent evaluation activity diminishes performer internalization and acceptance, will very frequent evaluation activity enhance internalization and acceptance? The answer would seem to be "no". Very frequent evaluation activity may also diminish internalization and acceptance. When evaluation activity is very frequent, it may interfere with the rational operation of the system. Evaluations performed prior to the completion of segments of task performance are likely to be perceived by performers as intrusive and as an interference with the completion of the allocated tasks. Moreover, immediate evaluations are just as likely to appear arbitrary if it seems that evaluators have not exercised careful judgement collecting adequate samples of performances and outcomes and formulating appraisals. Very frequent evaluations are thus less likely to lead to performer internalization and acceptance.

Considering the rational or operating effects of evaluation activities leads to the conclusion that there is some moderate frequency of evaluation activities likely to lead to optimum performer internalization or acceptance. The relationship between frequency of evaluation and performer acceptance may be described as curvilinear.

Considering the social or symbolic effects of evaluation activities leads to a similar conclusion. The rational or operating perspective on

evaluation assumes that tasks in organizations are fully defined and subjected to evaluation processes in a straightforward manner. The social or symbolic perspective argues that the evaluation process itself functions to socially and symbolically define tasks, the relevant performances and outcomes, the beginnings and ends of organizationally relevant activities. Evaluations help to define and highlight performances as relevant by providing an audience for them.

From the social perspective, very infrequent evaluations would appear to performers as sporadic and arbitrary, only rarely designating a bit of performance as relevant and important. Recipients of very infrequent evaluations might not even understand the significance of the evaluation activities. On the other hand, very frequent evaluations would appear to define everything as equally relevant and important. Because evaluation would be such a common activity, it would be useless as a device for defining and highlighting critical aspects of performance. Thus the social or symbolic perspective also suggests that optimum internalization or acceptability will occur when evaluations are moderately frequent.

While the precise meaning of "moderately" frequent evaluations leading to optimum internalization and acceptance cannot be specified at this point, it is possible to predict that the appropriate frequency for optimum internalization and acceptance will vary with the nature of the tasks being evaluated. It might at first seem that some "objective" task dimension such as task duration would affect the relationship between frequency of evaluation and performer internalization or acceptance. After all, it would be reasonable to assume that tasks that take longer to complete would require less frequent evaluation. However, tasks are socially defined and allocated in organizations. This means that in many cases there is no such thing as duration inherent to a task. Long tasks can often be broken up

into a series of shorter tasks. The task of building an automobile is a case in point. In modern automobile plants this task is broken into many tasks of shorter duration that are assigned to different performers. It is deemed appropriate to evaluate the performance of these shorter tasks throughout the production process; enter the quality control function.

Other tasks such as developing strong moral character in children are of long duration and do not seem amenable to division into smaller tasks of shorter duration. Tasks of this sort are typically thought of as more unpredictable; they often require the performer to monitor progress and adjust his or her activities in the middle of the performance. Evaluation is performed throughout this process, but the evaluation is self-evaluation by the performer. In fact, the self-evaluative act is the essence of the performance of unpredictable tasks. When evaluation by a superior takes place, the main focus of the superior's evaluation is the quality of the overall self-evaluation done by the performer. Unpredictable tasks thus defy external evaluation until they are completed. Much of teaching involves unpredictable tasks which require the teacher to act as a self-evaluator in the course of performing the task.

From the rational or operating perspective, it seems that unpredictable tasks should lower the frequency of evaluation associated with optimum acceptance. Performers who conceive of their tasks as unpredictable will find evaluations intrusive into the performance of the task at a much less frequent rate than performers who conceive of their tasks as predictable, for such evaluation may disrupt the process of self-evaluation.

This will also be the case from the social or symbolic perspective. For predictable tasks with few requirements for performers to adjust

activities during the course of performance, evaluation may be used to redefine the task into a series of tasks of shorter duration. Assembling an automobile becomes: a) assembling the chassis, b) assembling the engine...., each of which is simpler than the original task and each of which is amenable to more frequent evaluation. Thus there may be social agreement on the redefined tasks and evaluation may be the symbol which confirms the dimensions of the new tasks. For unpredictable tasks, it is impossible to obtain such social agreement, and more frequent evaluation becomes not a symbol of the redefined tasks, but a bizarre intrusion into the performance. Daily observance of a teacher's attempt to develop strong moral character in a student might be such a bizarre intrusion. Thus from the social perspective, more unpredictable tasks should require less frequent evaluation in order to achieve optimum acceptance or internalization.

The preceding discussion suggests the following proposition on the relationship between frequency of evaluation and performer internalization or acceptance of the evaluation process:

Proposition 1 - The more frequently performers experience evaluation activities, the more likely they will be to accept or internalize the evaluation process. However, after a certain point increasing the frequency of evaluation activities will diminish performer acceptance. The level of evaluation frequency corresponding with maximum performer acceptance decreases as task predictability decreases.

Some readers will recognize this first proposition as having a structure parallel to the Yerkes-Dodson Law (1908) relating aversive stimulation to learning.

Proposition 2 - Performer Influence and Internalization

The relationship between performer influence and internalization of the evaluation process may also be viewed from both the rational perspective and the social perspective. Once again, both perspectives lead to the same conclusions about the relationship.

Performers of tasks have the most intimate and direct knowledge of the work situation. They possess a knowledge beyond that of their supervisors who are at least one step removed from the actual work. An evaluation system that can be influenced by the performers of tasks should thus be more sophisticated and more appropriate to the particular tasks involved. From the rational perspective, such an evaluation system should function better than one that denies performer influence. For example, performers have knowledge of the work that places them in the best position to set standards for performance and output that are challenging without being frustrating, a fact which the countless time and motion studies of industrial engineers both attests to and ignores. Such studies typically treat workers as performers to be examined rather than colleagues with important information to share. Evaluation systems which permit some level of performer influence should thus be more appropriate and should, in turn, lead performers to internalize them more readily than those which deny any worker influence.

But once again, the question must be raised: If some performer influence is better than no performer influence, will a great deal of performer influence lead to a better evaluation system and greater internalization than one with moderate performer influence? The answer would appear to be "no". An evaluation system that is open to a great deal of performer influence may suffer from several problems. First, such a

system may not be stable enough to endure in a reliable form and may thus appear to be an arbitrary political artifact. Unless an evaluation system is consistent and reliable over a period of time, it is not likely to inspire confidence or to have an appreciable effect in leading to improved performance. If performers and supervisors are constantly renegotiating the evaluative norms, there may be little time left for getting on with the main tasks of the organization.

A second problem with evaluation systems that permit great performer influence has to do with the composition effects of multiple performers each exerting influence over the system. In such a situation there may not be a consensually agreed upon evaluation system in operation. Rather, each supervisor-subordinate dyad might evolve a unique and distinctive evaluative process. This would eventually lead to a breakdown in the very coordination and control processes evaluation systems are designed to accomplish. Thus extreme levels of performer influence over the evaluation process would impede, not enhance, the rational operation of the evaluation system in an organization. This, of course, suggests that the relationship between performer influence over the evaluation process and performer internalization or acceptance of that process is a curvilinear one with a moderate level of influence most likely to produce the greatest performer internalization of the evaluation process.

From the social perspective the relationship looks much the same. There have been enough studies of the effects of participation in programs and policies on individual involvement and commitment (See, for example, McLaughlin and Marsh, 1978), to conclude that systems that permit performer influence are more likely to secure acceptance and internalization than those which do not permit such influence. This would appear to be true even if the results of that influence do not lead to improvements in the

operation of the system. Above and beyond any improvements in the rational operation of the evaluation system, systems that permit some performer influence should be more successful in securing performer acceptance than those which permit no performer influence. Performers who exert influence on the shape of the evaluation process will be more likely to internalize the process since they participated in its creation.

From the social perspective, extremely high levels of performer influence should also diminish performer acceptance or internalization of the process. Evaluation systems that are influenced by the individual performer to an overwhelming degree lose their social character and approach the status of self-evaluations. While self-evaluations may at first appear to suffer from none of the problems that prevent organizational evaluation systems from being accepted by performers, these performer-determined evaluation processes, in fact, deprive performers of social affirmation of their own self-evaluations. Extremely high levels of performer influence over the evaluation process will thus lead to less acceptance than more moderate levels.

As was the case with frequency of evaluation activities, it is not possible to specify precisely the meaning of "moderate" performer influence over the evaluation process. However, once again, the appropriate degree of performer influence for optimum acceptance of the evaluation system would seem to vary with the nature of the tasks being evaluated. Once again, the critical task dimension is task predictability.

It has already been noted that self-evaluation is an integral aspect of the performance of unpredictable tasks. Accomplishing such unpredictable tasks as developing strong moral character in children requires performers who evaluate their own performance in the context of a

developing situation and modify their behavior based upon information acquired in the course of task performance. While there is nothing intrinsic to the nature of unpredictable tasks which would prevent supervisors from developing a complete understanding of the developing challenge of the task and the performance of the subordinate in a changing context, in practice the amount of time required to communicate this information to a supervisor would be prohibitively expensive for most organizations. In most cases only information on exceptional situations or aspects of performance can be passed upward to a supervisor. In such situations supervisors must depend upon subordinates for crucial perspectives on the task performance situation.

From the rational or operating perspective, unpredictable tasks will require greater performer influence in the evaluation process in order to maintain the quality of the evaluation process. For example, in the supervision of the performance of very unpredictable tasks, supervisors are more likely to apply criteria and standards inappropriately without first hand information from subordinates. This information may concern the applicability of certain criteria and standards to particular situations or the identification and selection of appropriate samples of performance information for particular situations. If performer influence over the evaluation process for unpredictable tasks is too limited, pre-determined evaluation activities may fail to meet the increased needs for adjustments in the application of criteria or for adequate information on performance.

From the social or symbolic perspective less predictable tasks will also suggest the need for greater performer influence over the evaluation process. Quite apart from the operational needs of the evaluation system, individual performers involved in unpredictable tasks will believe those tasks to be unique and will be less accepting of systems which do not

permit them to exercise influence over the process of evaluation.

The preceding discussion suggests the following proposition on the relationship between performer influence over the evaluation process and performer internalization or acceptance of the evaluation process:

Proposition 2 - The more influence performers have over evaluation activities, they more likely they will be to accept or internalize the evaluation process. However, after a certain point increasing the degree of performer influence will diminish performer acceptance. The level of performer influence corresponding with maximum performer acceptance increases as task predictability decreases.

3. Basic Features of the Six Studies Providing Evidence on the Two Propositions

The six studies that provide the major source of evidence related to the two propositions outlined above all focused on the impact of evaluation processes on teachers. The basic dimensions of the studies are presented in Table 1 below:

The first study in Table 1 (Dornbusch and Scott 1975; Thompson, 1971; Thompson, Dornbusch and Scott 1975) involved surveys of 131 teachers in six elementary schools in a small school district. Approximately 85% of the

Table 1
Basic Features of Six Studies Using the Theory of Evaluation and Authority

Studies (Sources)	Number of: Schools	Teachers	Independent Variables	Dependent Variables
1) Questionnaire study of public elementary school teachers and interviews with their principals (Thompson 1971; Thompson, Dornbusch and Scott, 1975)	6	131	Frequency of Communicated Evaluations	Satisfaction with Evaluations Helpfulness of Evaluations
2) Comparative study of public school teachers and hospital nurses (Marram, 1972; Marram, Dornbusch, and Scott, 1972)	15	244	Frequency of Observations of Performance/ Outcomes	Soundness of Evaluations Importance of Evaluations
3) Comparative study of teachers in public schools and alternative schools (McCauley, 1971; McCauley, Dornbusch, and Scott, 1972)	29	200	Frequency of Communicated Evaluations Frequency of Observations of Performance	Satisfaction with Evaluations Soundness of Evaluations
4) Intensive interview study of public school teachers (Natriello and Rowe, 1981)	1	18	Frequency of Observations of Performance Frequency of Communicated Evaluations Influence Over Criteria-Setting	Satisfaction with Evaluations Helpfulness of Evaluations Preferences for Evaluation to Influence Rewards and Penalties
5) Questionnaire study of urban middle school teachers (deCharms and Natriello, 1981; Natriello, forthcoming)	6	182	Frequency of 5 Stages of Evaluation Influence over 6 Stages of Evaluation	Soundness of Evaluations Leverage
6) Interview study of suburban elementary school teachers working under a merit pay system (Natriello and Cohn, 1983)	1	23	Frequency of 5 Stages of Evaluation Influence over 6 Stages of Evaluation	Soundness of Evaluations Helpfulness of Evaluations Preferences for Influence on Rewards and Penalties Leverage

teachers in the district participated in the study, which was conducted under the auspices of a committee of teachers designed to review the arrangements for the evaluation of teachers and recommend improvements. The study provided data on the effects of both frequency of evaluation and teacher influence over the evaluation process. Dependent variables included teacher satisfaction and teacher perceptions that the evaluations they received were helpful in improving their performance.

The second study in Table 1 (Dornbusch and Scott, 1975; Marram, 1972; Marram, Dornbusch and Scott, 1972) involved surveys of hospital nurses and public elementary school teachers. Responses from 244 teachers in fifteen elementary schools in a single school district provided data on several measures of the frequency of evaluation, also referred to as task visibility. Dependent variables included teacher perceptions that the evaluation system was soundly based and that evaluations were important.

The third study presented in Table 1 involved surveys of teachers in public schools and teachers in alternative schools. This study (Dornbusch and Scott 1975; McCauley, 1971; McCauley, Dornbusch, and Scott, 1972) involved surveys of teachers in public schools and teachers in alternative schools. The 100 public school teachers were drawn from five schools in two school districts, while the 100 alternative school teachers came from twenty-four San Francisco Bay Area alternative schools. Teacher responses provided data on the frequency of evaluation activity. Dependent variables included teacher satisfaction with the evaluation process.

The fourth study (Natriello and Rowe, 1981) included surveys and interviews with all of the teachers in a single elementary school in a suburban school district. In this intensive case study, 18 teachers provided data on the frequency of evaluation and teacher influence over

aspects of the evaluation process. Dependent variables included satisfaction with the evaluation process, teacher perceptions of the helfulness of the evaluation process, and teacher preferences for the influence of evaluations over the distribution of organizational rewards and penalties.

The fifth study (deCharms and Natriello, 1981; Natriello, forthcoming) was conducted in six middle schools in a major urban area in the midwest. Surveys of 182 teachers or 97% of the teachers in the six schools provided data on the frequency of evaluation activities and teacher influence over the evaluation process. Dependent variables included teacher perceptions of the soundness of the evaluation system and teacher leverage over teaching tasks.

The sixth study (Natriello and Cohn, 1983) was conducted in a single elementary school in a sururban school district with a merit pay program attached to the evaluation system. Survey responses from 23 teachers provided data on the frequency of evaluation activity and teacher influence over evaluation activities. Dependent variables included teacher perceptions of the helpfulness of evaluations, teacher preferences for the influence of evaluations on organizational rewards and penalties, teacher perceptions of the soundness of the evaluation system, and teacher leverage over teaching tasks.

4. Frequency of Evaluation, Teacher Influence Over Evaluations, and Indicators of Internalization in Schools

Frequency of Evaluation

Studies of processes for the evaluation of teachers suggest that evaluation is a relatively infrequent event. Reporting on the results of an HEW study, Dreeben (1970) noted that, on the average, 55% of

principals said they did not have enough time to conduct an accurate evaluation of classroom teachers. In the same study over one-fourth of the classroom teachers reported no classroom visitations by any supervisor during the first half of the 1962-1963 school year. Further, almost one-fifth of the probationary teachers, a group likely to be more closely supervised, received no visits during the same period.

The six studies guided by the theory of evaluation and authority seem to confirm this pattern. In the study by Thompson (1971) teachers were asked how frequently their principals communicated evaluations of their performance to them. Thompson (1971) reported that the majority of teachers indicated that they received communications of evaluations from their principal "seldom" or less frequently for each of four defined teaching tasks. The proportions of teachers reporting they received evaluations from their principal "seldom" or less frequently were .56, .58, .62, and .62 for the tasks of teaching subject matter, character development, maintaining control and record keeping, respectively.

The study conducted by Marram (1972) also asked teachers about the frequency of evaluation activities. In this study teachers were asked how frequently their principal observed their performance for each teaching task and how frequently their principal observed the outcomes of their performance of each teaching task. Using a scale of responses consisting of: 1 - "very frequently", 2 - "frequently", 3 - "fairly often", 4 - "occasionally", 5 - "seldom", 6 - "almost never", and 7 - "never", Marram (1972) calculated the median responses for each teaching task for each survey item. The median responses for the item pertaining to the frequency of observations of performances were 3.6, 3.6, 3.4, and 4.1 for the tasks of teaching subject matter, character development, maintaining control, and record keeping, respectively. The median responses for the item pertaining

to observations of outcomes were 3.3, 3.3, 3.2, and 3.9 for the four teaching tasks respectively.

In the study of teachers in public schools and alternative free schools, McCauley (1971) asked teachers to indicate how frequently they learned of their principal's evaluation of their performance on each of the four teaching tasks. Using the same scale of responses used by Marram, McCauley reported median scores from the public school teachers of 4.0, 4.2, 4.2, and 4.9 for the tasks of teaching subject matter, character development, maintaining control, and record keeping, respectively. The median scores on the same item for the alternative school teachers were 3.6, 3.9, 4.1, and 6.2 for the four teaching tasks, respectively.

An in-depth study of a single elementary school conducted by Natriello and Rowe (1981) revealed a similar pattern. Following the task specific approach of Dornbusch and Scott, they asked teachers to report on evaluation of their teaching in terms of four teaching tasks: teaching subject matter, character development, maintaining control, and record keeping. The proportions of teachers reporting that the principal observed their performance seldom or less frequently were .22, .22, .22, and .11 for the tasks of teaching subject matter, character development, maintaining control, and record keeping, respectively. The same proportions of teachers reported that the principal observed the outcomes of their performances seldom or less frequently. In addition, for each task, the proportion of teachers who did not know how frequently their principal observed their performance or the outcomes of their performance was .22.

The pattern of infrequent evaluation of teaching holds when teachers are asked about the frequency with which they receive feedback from their principal on their performance of the teaching tasks. The proportions of

teachers reporting that they received feedback seldom or less frequently were .56, .44, .68, and .78 for the tasks of teaching subject matter, character development, maintaining control, and record keeping respectively.

When asked how often their principal identified an area in which they needed to improve for each of their teaching tasks, all of the teachers in the study reported that this happened seldom or less frequently for each teaching task. Teachers were then asked how often their principal provided them with help to improve their performance on those occasions when an area needing improvement was identified. The proportions of teachers reporting that their principal did this seldom or less frequently were .63, .50, .38, and .75 for the tasks of teaching subject matter, character development, maintaining control, and record keeping, respectively.

This case study of a single elementary school included in-depth interviews with the teachers. During the interviews teachers talked about their autonomy in the school as it related to the evaluation process. Their comments were characterized by a tone of ambivalance. On the one hand, over three-fourths of the teachers reported that they had total freedom on the tasks of teaching subject matter, and character development. Slightly fewer than three-fourths reported total freedom on the task of maintaining control, and still fewer reported having total freedom on the task of record keeping. The teachers seemed generally satisfied with their autonomy in the school.

However, most of the teachers also reported that this feedom had its price. The lack of evaluative activity resulted in feelings of uncertainty and confusion as they went about their tasks. A teacher who had worked under this principal for a number of years explained that:

I learned by trial and error. I learned a long time ago I won't

get a go-ahead from him. If we do something 'good' everything is fine, if not, your're out on a limb. I don't know what he wants. I do it my way and its acceptable but new teachers don't know that. I can't tell them--I might be wrong. I don't know what he wants--I just worked it out.

Another teacher expressed the isolation felt by many teachers in the school:

I told him (the principal) it's so lonely down there by myself when you don't give me feedback. Come to my room and see. He says, "I know you do a good job, etc." I don't buy it. He needs to be in my room.

These comments from teachers suggest only some of the consequences of the low levels of evaluation activity found in study after study in schools.

In the study of teachers in six inner-city middle schools, Natriello (forthcoming) asked teachers to report on the frequency with which they experienced activities in five of the six stages of the evaluation model. Teachers were not asked to report on the frequency of appraisal since they would be unlikely to know how often their principal engaged in this process.

For each stage of the evaluation process, substantial proportions of teachers reported that activities occurred seldom or less frequently. These proportions are represented in the graphs in Figure 2, section A. The proportions of teachers reporting that they seldom or less frequently received task assignments from their principal telling them the goals for each task were .54, .48, .47, and .39, respectively, for the tasks of teaching subject matter, character development, maintaining control, and record keeping. The proportions reporting that they seldom or less

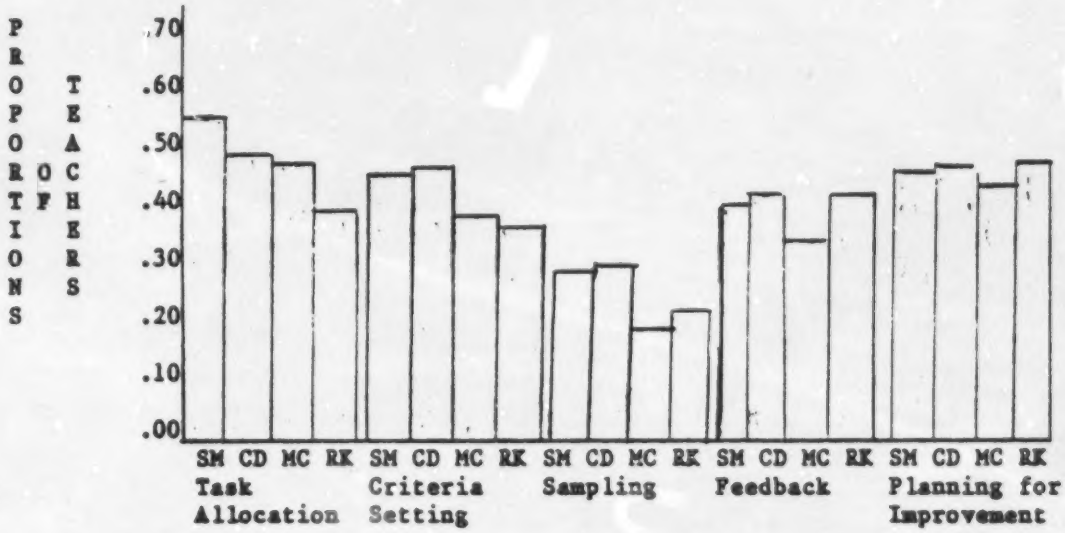
frequently learned of the criteria and standards for performance on each task were .45, .46, .38, and .35 for the tasks of teaching subject matter, character development, maintaining control, and record keeping, respectively. The proportions reporting that their principal seldom or less frequently observed aspects of their performance were .28, .29, .18, and .21 for the four teaching tasks, while the proportions reporting that they seldom or less frequently received feedback on each task were .40, .41, .33, and .41 for the tasks of teaching subject matter, character development, maintaining control, and record keeping, respectively. Finally, the proportions reporting that their principal worked with them to plan ways to improve their performance seldom or less frequently were .44, .45, .42, and .47 for the four teaching tasks.

But evaluation does not have to be so infrequent. The study by Natriello and Cohn (1983) of evaluation practices in a single elementary school in a suburban district with a long standing merit pay program revealed more frequent activity in most of the five stages of the evaluation process. These proportions are represented in the graph in Figure 2, section B. In that study the proportions of teachers reporting task assignments in which the principal communicated their goals for their teaching tasks seldom or less frequently were .35, .55, .60, and .50 for the tasks of teaching subject matter, character development, maintaining control, and record keeping, respectively. The proportions reporting that they learned of the criteria and standards used to evaluate their

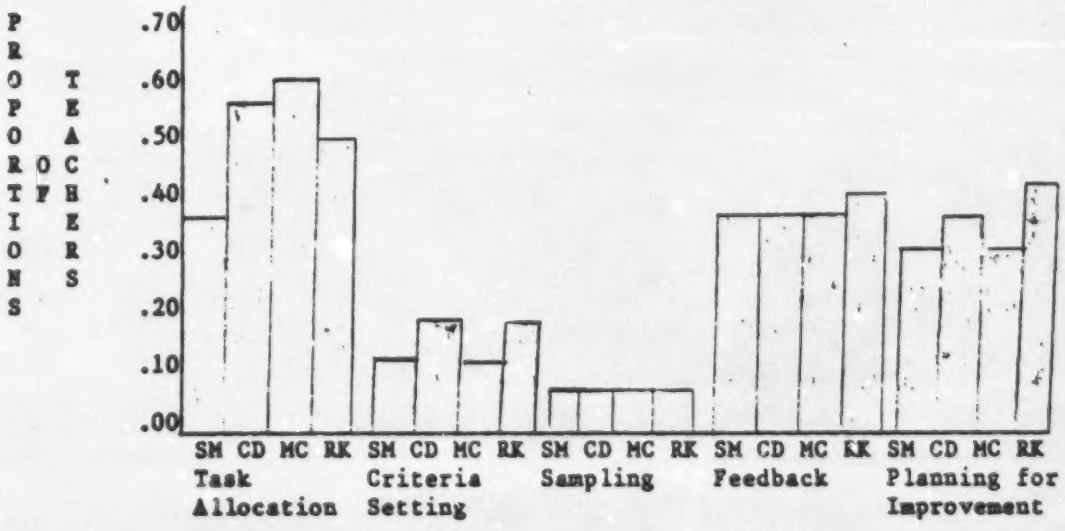
Figure 2

Proportions of Teachers in The Urban Middle Schools Study and the Merit Pay District Study Reporting Experiencing the Evaluation Activities "Seldom" or Less Frequently, by Teaching Task

A. Urban Middle Schools Study



B. Merit Pay District Study



performance only seldom or less frequently were .11, .17, .11, and .17, for the four teaching tasks. The proportions of teachers reporting that their principal observed their performance only seldom or less frequently were .05, .05, .05, and .05 for the four teaching tasks.

The proportions of teachers who reported receiving feedback on their performance on each task only seldom or less frequently were .35, .35, .35, and .40 for the tasks of teaching subject matter, character development, maintaining control, and record keeping, respectively. Finally, the proportions of teachers who reported that their principal worked with them to plan ways to improve their performance only seldom or less frequently were .30, .35, .30, and .42 for the four teaching tasks.

This series of studies reveals that the activities identified in the model of the evaluation process are relatively infrequent events for many teachers and that the frequency of any one activity may differ from that of the other activities. For example, in a district with more frequent evaluation activity, such as the suburban school district studied by Natriello and Cohn (1983), although very small proportions of teachers reported that criteria setting and sampling were infrequent events, over one-third of the teachers reported that they seldom or less frequently received feedback on each of their teaching tasks.

Teacher Influence Over the Evaluation Process

Four of the six studies collected data on teacher perceptions of their influence over various stages of the evaluation process. Thompson (1971) asked teachers to report on their influence over the criteria setting process and over the sampling process. In that study the proportions of teachers reporting that they were only "slightly influential" or "not at all influential" in affecting the criteria setting process were .60, .61.

development, maintaining control, and record keeping, respectively. As for their influence over the selection of information used for evalution purposes, the proportions of teachers reporting that they were only "slightly influential" or "not at all influential" were .69, .70, .58, and .75, for the four teaching tasks, respectively.

In the intensive study of the single elementary school teachers were asked how much input they had in the selection of criteria used to determine their evaluations. The proportions of teachers reporting little or no input were .44, .33, .33, and .56 for the tasks of teaching subject matter, character development, maintaining control, and record keeping, respectively. They were also asked to report on whether they had any input into two task allocation decisions, the decision about which grade level they would be assigned and the decision about which students would be assigned to their class. Eighty percent of the teachers reported that they had no input into the decision as to which grade level they would be assigned, and .38 reported that they had no input in determining which students would be assigned to their class.

In the study of six inner-city middle schools, Natriello (forthcoming) asked teachers to report on their influence over the six stages of the evaluation process. Teachers reported having substantial influence in determining task allocations, and the goals of their teaching. These results are portrayed in Figure 3, section A. The proportions of teachers reporting having only slight influence or no influence at all over task allocations were .11, .12, .11, and .31, respectively, for the tasks of teaching subject matter, character development, maintaining control, and record keeping.

When it came to the second stage of the evaluation process teachers

were less likely to report having influence. The proportions of teachers reporting little or no influence over this aspect of the evaluation process were .47, .49, .51, and .53 for the tasks teaching subject matter, character development, maintaining control, and record keeping, respectively. A similar pattern was revealed for teacher influence over the sampling process. The proportions of teachers reporting little or no influence over sampling were .53, .54, .52, and .55 for the four teaching tasks.

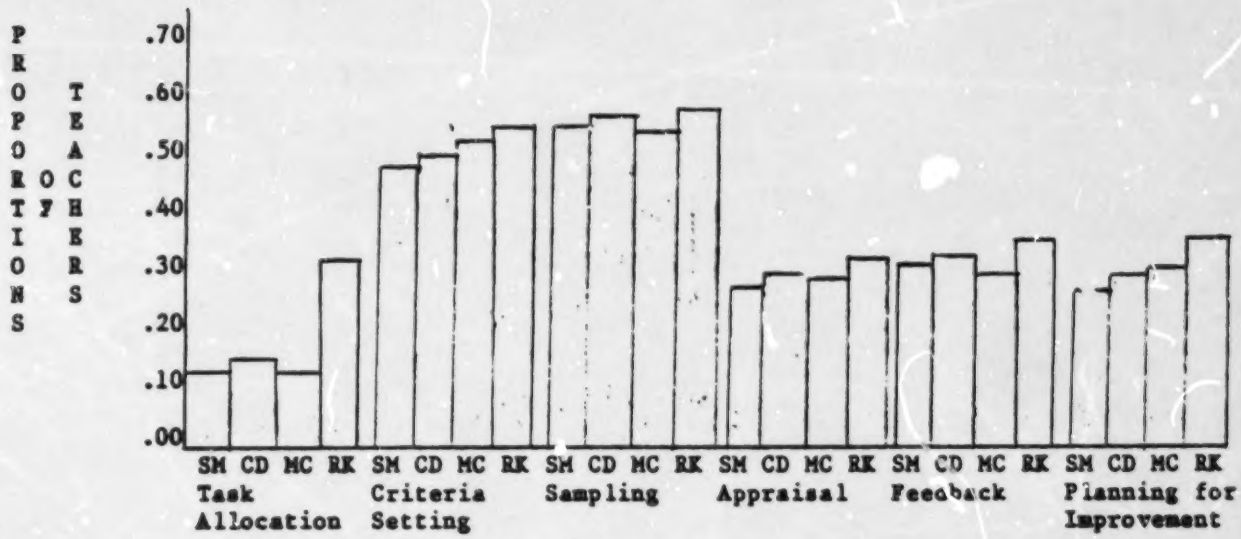
More teachers reported having influence over the last three stages of the evaluation process. The proportions of teachers reporting little or no influence over the appraisal process were .26, .29, .28, and .32 for the tasks of teaching subject matter, character development, maintaining control, and record keeping, respectively. For influence on the feedback process the corresponding proportions were .31, .32, .29, and .35. For influence on the improvement process the corresponding proportions were .26, .29, .30, and .35.

The study of the single elementary school in the suburban district with the long-standing merit pay system revealed that teachers in this school were able to exercise more influence over the evaluation process. At every stage of the evaluation process teachers were less likely to report that they were only slightly or not at all influential. The results of the study are portrayed in Figure 3, section B. The proportions of

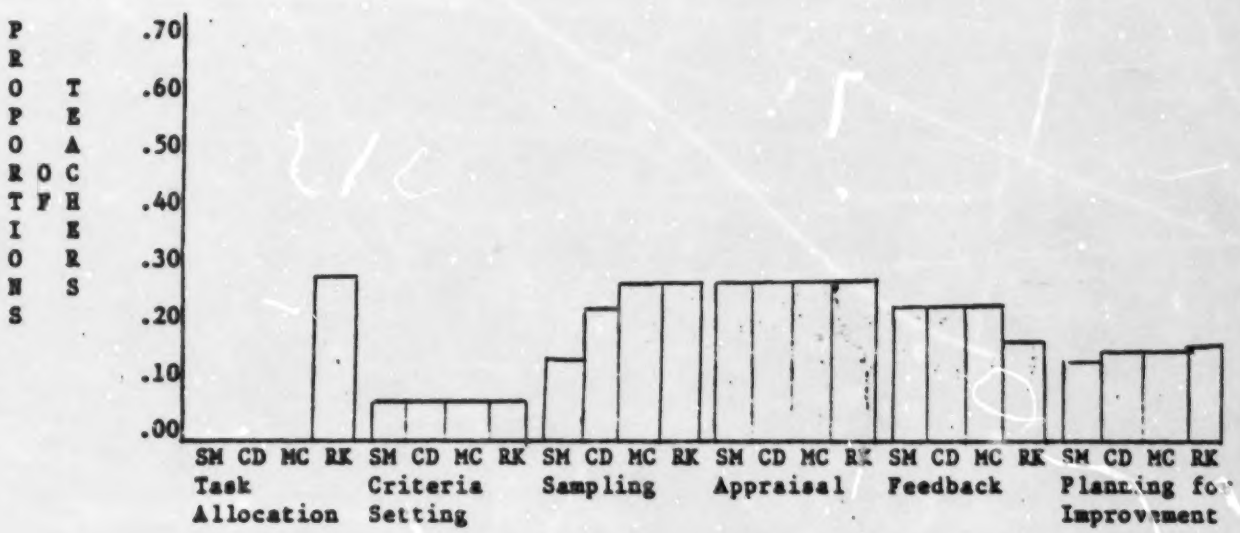
Figure 3

Proportions of Teachers in The Urban Middle Schools Study and the Merit Pay District Study Reporting That They are Only "Slightly Influential" or "Not at All Influential" in Affecting the Stages of the Evaluation Process, by Teaching Task

A. Urban Middle Schools Study



B. Merit Pay District Study



teachers reporting that they were slightly or not at all influential in determining their goals for their teaching tasks were .00, .00, .00, and .25 for the tasks of teaching subject matter, character development, maintaining control, and record keeping respectively. The corresponding proportions for influence over the criteria setting process were .05, .05, .05, and .05 for the four teaching tasks.

The proportions of teachers reporting that they were slightly or not at all influential in regard to the sampling process were .10, .21, .26, and .26 for the tasks of teaching subject matter, character development, maintaining control, and record keeping, respectively. The corresponding proportions for the influence of teachers on the appraisal process were .26, .26, .26, and .26.

The proportions of teachers reporting that they were slightly or not at all influential in affecting the feedback process were .21, .21, .21, and .16, respectively, for the tasks of teaching subject matter, character development, maintaining control, and record keeping. The corresponding proportions for the process of planning for improvement were .10, .11, .11, and .15.

While teachers in some schools seem to exercise less influence over the evaluation process than they might prefer, evidence from the merit pay district where evaluation activities are more frequent suggests that even in such developed systems teachers can exercise considerable influence over the evaluation process. The important point for the present analysis is that increased attention to evaluation activity and greater evaluation activity in a school does not have to diminish teacher influence.

Indicators of Teacher Internalization or Acceptance of the Evaluation System

The six studies reviewed here used a variety of measures that might be considered indicators of teacher internalization or acceptance of the evaluation system. Although each indicator is in some way unique, all of the indicators provide some understanding of Duckworth's (1981) concept of internalization.

Three of the studies (Thompson, 1971; McCauley, 1971; and Natriello and Rowe, 1981) employed a measure of teacher satisfaction with the process by which their performance was evaluated. In these studies teachers were asked to simply indicate how satisfied they were with the way their performance was evaluated.

Four studies (Thompson, 1971; Natriello and Rowe, 1981; Natriello, forthcoming; Natriello and Cohn, 1983) collected data on teacher perceptions of the helpfulness or usefulness of the evaluations they received. In the studies by Thompson, 1971, and Natriello and Rowe (1981) teachers were asked to rate the helpfulness of the evaluations they received. In the studies by Natriello (forthcoming) and Natriello and Cohn, (1983) teachers were asked to indicate how frequently the information provided by their evaluations actually enabled them to work toward improving their performance.

One study (Marram, 1972) included a question which asked teachers how important the evaluations they received were to them. Assuming that those teachers who believe evaluations are more important have internalized the system allows this measure to be treated as an indicator of internalization or acceptance.

In three studies (Marram, 1972; Natriello and Rowe, 1981; and Natriello and Cohn, 1983) teachers were asked to report how much influence

they preferred evaluations of their performance to have on the organizational rewards and penalties they received. This question really probed the extent to which teachers wanted evaluations of their performance to "count", i.e. to have real consequences for them above and beyond the communicated evaluation.

Three studies (Marram, 1972; Natriello, forthcoming; Natriello and Cohn, 1983) examined teacher reports on the extent to which the evaluation system was soundly based. The studies by Marram, (1972) and Natriello and Cohn, 1983) asked teachers to indicate how soundly based their principal's evaluations of their performance were, while the study by Natriello, (forthcoming) asked teachers to report how frequently their principal's evaluations were soundly based.

In addition to these items which left the interpretation of "soundly based" up to the individual teachers, the studies by Natriello (forthcoming) and Natriello and Cohn (1983) also contained items based on the formal definition of "soundly based" contained in the theory of evaluation and authority. The theory of evaluation and authority specifies the following definition for soundly based evaluations:

> A participant considers evaluations <u>soundly based</u> to the extent that he or she believes that (a) the quality of performances or outcomes as judged by the participant is affected by the performer's effort, and (b) performances or outcomes considered better by the participant receive higher evaluations. (Dornbusch and Scott, 1975:343)

The definition thus contains two elements. The first is that performers consider evaluation soundly based to the extent that the performer's effort

is perceived to affect the evaluator's judgement. The second is that evaluators give higher evaluations to the performances or outcomes that performers consider to be better. Soundly based evaluations entail two relationships--the relationship between performer effort and performances or outcomes and the relationship between performances/outcomes and evaluations. In soundly based systems of evaluations higher levels of performer effort will result in better evaluations. In the study by Natriello and Cohn (1983) teachers were asked how often high levels of effort resulted in good evaluations and how often high levels of performance resulted in good evaluations. These items were also used as measures of the extent to which the evaluation system was soundly based.

Two studies (Natriello, forthcoming, Natriello and Cohn, 1983) contained indicators of a concept termed "leverage". Leverage refers to the relationship between the effort put forth by a performer and the outcomes that result from that effort. Leverage may be described as the ratio between outcome and effort and may range from 1/0 to 0/1. 1/0 represents the case where for 0 units of effort the performer realizes 1 unit of output. 0/1 represents the case where the performer realizes 0 units of output for every 1 unit of effort. Of course, 1/1 represents the case where for every unit of effort the performer realizes 1 unit of outcome.

The two studies contained items which provided data for two measures of teacher leverage over their tasks. One measure, called the Teacher-Assessment of Leverage, required teachers to note which combination of effort ratings and effectiveness ratings best described their situation during the past year. Response categories consisted of nine possible combinations of three ratings of effort (high, medium, low) and three ratings of effectiveness (high, medium, low). Leverage was scored as

greater than or equal to 1 whenever the effectiveness rating was equal to or greater than the effort rating and as less than 1 whenever the effectiveness rating was less than the effort rating.

A second measure, called the Composite Assessment of Leverage measure, was created from teacher responses to separate items which asked them to report on their current levels of effort and effectiveness compared to levels earlier in their teaching careers. If we keep in mind Duckworth's condition that evaluation systems should improve teacher efficacy, then it is reasonable to consider teacher leverage another indicator of teacher internalization or acceptance of the evaluation process.

The studies considered here together have relied on six dependent variables as results of particular evaluation processes: 1) teacher satisfaction, 2) teacher perceptions that the evaluations are useful in helping them to improve their performance, 3) teacher perceptions that evaluations are important, 4) teacher preferences for evaluations to have an impact on organizational rewards and penalties, 5) teacher perceptions that evaluations are soundly based, and 6) teacher leverage over teaching tasks. In examining the impact of frequency of evaluation activities and teacher influence over evaluations on teacher acceptance of the evaluation system, each of these six indicators will be considered.

5. The Impact of Frequency of Evaluation and Teacher Influence over Evaluation Processes on Teacher Acceptance of the Evaluation System

The six studies reviewed here permit an examination of the impact of the frequency of evaluation and teacher influence over the evaluation process on a range of indicators of teacher acceptance of the evaluation system. In this presentation only evidence bearing on the positive relationship

between frequency of evaluation and teacher acceptance of the evaluation process and bearing on the positive relationship between teacher influence over the evaluation process and teacher acceptance of the evaluation process will be considered. Of course, the original hypotheses of curvilinear relationships necessitated an examination of the upper end of the distributions for frequency of evaluation and teacher influence over the evaluation process. This was done first by direct inspection of the plotted data points and then by dividing the distributions for each of these variables and examining the relationships between them and the indicators of teacher acceptance or internalization along both the bottom half and the top half of the range. These analyses revealed no substantial negative effects from either variable. The complete analyses for the urban middle schools study are presented in Natriello and Dornbusch (forthcoming). To simplify the presentation, only the relationships between these variables and the indicators of acceptance along the entire range are reported here. Table 2 presents the results of analyses, using data from all six studies, of the relationship between various measures of the frequency of evaluation activities and the indicators of teacher acceptance.

The statistical relationships between variables are expressed in gammas. Gamma is a nonparametric (making no assumptions about the underlying distribution) measure of the relationship between two variables. Gamma may be interpreted as a measure of the proportionate reduction in error; it measures the extent to which knowing the value of one variable for one case reduces error in predicting the value of a second variable for that same case. Gamma ranges from +1 to -1. A positive gamma means that the order on one variable is similar to the order on the other variable; a negative gamma indicates that the order on one variable is inversely related to the order on the other.

Table 2
Relationship Between Frequency of Evaluation Activities and Indicators of Teacher Acceptance of the Evaluation System

Frequency Measure	Acceptance Indicator	Gammas by Teaching Task				Source Study
		Subj. Mat.	Char. Dev.	Main. Cont.	Rec. Keep.	
Communicated Evaluations	Satisfaction	.68	.68	.68	.55	1
Principal Observations of Performance	Satisfaction	1.00	1.00	1.00	.71	4
Communicated Evaluations	Satisfaction	1.00	.71	1.00	1.00	4
Communicated Evaluations	Helpfulness	.84	.79	.81	.70	1
Observations of Performance	Helpfulness	.43	.25	-1.00	-1.00	4
Communicated Evaluations	Helpfulness	.43	.25	-.25	-1.00	4
Summary of Five Stages	Information Helpful	.56	.72	.52	.80	5
Summary of Five Stages	Principal Helpful	.48	.80	.66	.75	5
Summary of Five Stages	Helpfulness	-.33	1.00	-.09	.50	6
Observations of Performance	Importance	.19	.55	.39	.63	2
Observations of Outcomes	Importance	.30	.58	.45	.58	2
Observations of Performance	Prefer Related to Rewards and Penalties	.50	.50	.00	.50	4
Communicated Evaluations	Prefer Related to Rewards and Penalties	.14	-.50	.00	1.00	4
Summary of Five Stages	Prefer Related to Rewards and Penalties	1.00	1.00	1.00	1.00	6

Observations of Performance	Soundly Based	.76	.78	.76	.52	2
Observations of Outcomes	Soundly Based	.60	.70	.64	.61	2
Summary of Five Stages	Soundly Based	.51	.61	.64	.49	5
Summary of Five Stages	Soundly Based	.92	.70	-.14	.58	6
Summary of Five Stages	Good Evaluations if Do Well	1.00	1.00	-1.00	.44	6
Summary of Five Stages	Good Evaluations if Try Hard	.25	1.00	-1.00	.56	6
Summary of Five Stages	Teacher Assessment of Leverage	.51	.29	.52	.24	5
Summary of Five Stages	Composite Assessment of Leverage	.20	.39	.33	.39	5
Summary of Five Stages	Observed Leverage	.28	.26	.23	.34	5
Summary of Five Stages	Teacher Assessment of Leverage	.40	1.00	.20	-.07	6
Summary of Five Stages	Composite Assessment of Leverage	-.18	-.20	.00	-.66	6

Two of the six studies (Natriello and Rowe, 1981; Natriello and Cohn, 1983) were each conducted in a single elementary school and thus contain data on fewer than 30 teachers. While these data should be examined cautiously, they are included because they contribute to the general pattern of findings in the four larger studies.

Considering the results from all six studies, there are 100 possible instances to examine the relationship between the measures of frequency of evaluation and the indicators of teacher acceptance of the evaluation system. In 86 of these, the relationship is positive. In only 14 cases is the relationship negative. Considering only the four larger studies, there are 48 possible instances to examine the relationship. The relationship is positive in all 48 of these.

Thus the data from the six studies which included measures of the frequency of evaluation activities and teacher acceptance of the evaluation process overwhelmingly demonstrate that there is a positive relationship between frequency and teacher acceptance. Additional analyses (e.g. Natriello, forthcoming) examining only the high end of the distribution for the frequency measures as they relate to teacher acceptance revealed no evidence of the hypothesized curvilinear relationship.

Table 3 presents the results of analyses of the relationship between teacher influence over the evaluation process and teacher acceptance of the evaluation process. Considering the data from all six studies, there are 64 instances to examine the relationship. In 60 of these instances the relationship is positive, while in only 4 is the relationship negative. Considering only the data from the four larger studies, there are 24 instances to examine the relationship between teacher influence and teacher acceptance. In 23 of these instances the relationship is positive, while in only 1 instance is the relationship negative.

Table 3

Relationship Between Influence Over the Evaluation Process and Indicators of Teacher Acceptance of the Evaluation System

Influence Measure	Acceptance Indicator	Gammas by Teaching Task: Subj. Mat.	Char. Dev.	Main. Cont.	Rec. Keep.	Source Study
Influence Over Criteria-Setting	Satisfaction	.33	.71	.71	-.11	4
Influence Over Criteria-Setting	Helpfulness	.43	.25	.25	.20	4
Summary of Six Stages	Helpfulness	.33	.14	.82	.39	6
Summary of Six Stages	Information	.46	.53	.47	.51	5
Summary of Six Stages	Principal Helpful	.56	.63	.60	.61	5
Influence Over Criteria-Setting	Prefer Related to Rewards and Penalties	.14	.50	.00	.17	5
Summary of Six Stages	Prefer Related to Rewards and Penalties	1.00	.65	.22	.86	6
Summary of Six Stages	Soundly Based	.51	.61	.64	.40	5
Summary of Six Stages	Soundly Based	.86	.74	.64	.56	6
Summary of Six Stages	Good Evaluations if Do Well	1.00	1.00	.45	.66	6
Summary of Six Stages	Good Evaluations if Try Hard	.75	.76	.58	.74	6
Summary of Six Stages	Teacher Assessment of Leverage	.45	.27	.40	.25	5
Summary of Six Stages	Composite Assessment of Leverage	.21	.24	.34	.00	5

Summary of Six Stages	Observed Leverage	-.04	.09	.14	.10	5
Summary of Six Stages	Teacher Assessment of Leverage	.00	.35	.45	.08	6
Summary of Six Stages	Composite Assessment of Leverage	-.27	.50	.64	-1.00	6

The data from the six studies overwhelmingly demonstrate that the relationship between teacher influence over the evaluation process and teacher acceptance of the evaluation process is positive. Additional analyses at the high end of the influence range reveal no evidence of the hypothesized curvilinear relationship.

These positive relationships reported above persist when other relevant variables are controlled. The most relevant control variables in the context of the present discussion are, of course, frequency of evaluation and teacher influence over the evaluation process. Accordingly, Tables 4 and 5 present the analyses for the data from the urban middle schools study bearing on the relationships discussed above, this time using teacher influence as a control in examining the relationship between frequency of evaluation and teacher acceptance and using frequency of evaluation as a control in examining the relationship between teacher influence and teacher acceptance.

As these tables demonstrate, the relationships between frequency and teacher acceptance and between influence and teacher acceptance remain positive in the great majority of the cases even when controlling for the effects of the other independent variable.

As the data from the six studies show, the frequency of evaluation activities and teacher influence over the evaluation activities are each positively related to teacher acceptance of the evaluation process. The general low levels of frequency of evaluation and teacher influence over evaluation in these studies do not permit us to reject the original curvilinear hypotheses. This element of the formal propositions along with the impact on these relationships of the nature of the task, await further examination.

Table 4

Relationship Between the Frequency of Evaluation Activity and Indicators of Teacher Acceptance of the Evaluation System Under Conditions of Low and High Teacher Influence, in the Urban Middle Schools Study

Acceptance Indicator	Influence Condition	Gammas by Teaching Task			
		Subj. Mat.	Char. Dev.	Main. Cont.	Rec. Keep.
Information Helpful	High	.60	.47	.34	.51
	Low	.79	.60	.72	.68
Principal Helpful	High	.76	.75	.79	.73
	Low	1.00	.91	.93	.70
Soundly Based	High	.75	.81	.74	.71
	Low	.72	.77	.81	.96
Teacher Assessment of Leverage	High	.62	.36	.39	.16
	Low	-.18	-.06	.30	.23
Composite Assessment of Leverage	High	.26	.20	.23	.43
	Low	-.40	.14	.49	.59

Table 5

Relationship Between Teacher Influence Over the Evaluation Process and Indicators of Teacher Acceptance of the Evaluation System Under Conditions of High and Low Frequency of Evaluation Activities, in the Urban Middle Schools Study

Acceptance Indicator	Frequency Condition	Gammas by Teaching Task			
		Subj. Mat.	Char. Dev.	Main. Cont.	Rec. Keep.
Information Helpful	High	.09	.33	-.10	.23
	Low	.45	.48	.42	.45
Principal Helpful	High	.50	.39	.27	.65
	Low	1.00	.75	.70	.61
Soundly Based	High	.33	.47	.36	-.56
	Low	.27	.40	.49	.39
Teacher Assessment of Leverage	High	.75	.33	.39	.04
	Low	.07	-.10	.30	.12
Composite Assessment of Leverage	High	.46	.26	.15	-.30
	Low	-.18	.19	.43	-.09

6. A Brief Examination of these Propositions Applied to the Evaluation of Students

The current discussion has focused exclusively on the evaluation of teachers. Since the theory of evaluation and authority applies to evaluation processes in general, a word is in order about the evaluation of another prominent group of performers in schools, students. An extensive examination of evaluation processes as they affect students is contained in Natriello and Dornbusch (in press) which presents a series of studies using the theory to examine the evaluation of students in secondary schools. These studies clearly demonstrate that more frequent evaluation of students results in greater student effort on school tasks. Moreover, these studies also show that those groups of students who traditionally do less well in school, minority students, receive less challenging evaluations than majority students.

Although the studies of the evaluation of students based on the theory of evaluation and authority conducted thus far provide no evidence of the negative impact of very frequent evaluation, given the role of the teacher as a primary evaluator of student performance and the physical proximity of teachers and students in the classroom/worksite, we may anticipate more instances of overly frequent evaluation of students than of teachers.

The matter of student influence over the evaluation process has not been explicitly considered in studies based on the theory of evaluation and authority. At first look, it may appear as if performer influence over the evaluation process is less pertinent to students who are not adults. After all, adolescents and children must constantly deal with situations where they have less influences than adults in the same situation.

Nevertheless, it seems likely that student influence over evaluation activities will lead to greater acceptance of the evaluation process. Such influence might be expressed in classrooms in perfectly reasonable ways.

Students are often able to influence teacher decisions regarding samples of performance used for evaluations. Some teachers even adopt policies that enhance student influence over the evaluation process. For example, teachers who institute policies permitting extra-credit work are extending to students opportunities to influence the sampling process. At this point there is little reason to believe that student influence over the evaluation activities will not have the same impact on student acceptance of the evaluation process as teacher influence has on teacher acceptance.

7. Directions for Further Research and Theory Construction

As demonstrated by the six studies of the evaluation of teachers, evaluation processes may differ both in their dimensions and in their effects on performers. These differences appear both for individual performers within a single organization and for different systems in different organizations. In this paper the emphasis has been on two dimensions of evaluation systems, the frequency of evaluation and the influence of performers, those being evaluated, over the evaluation process. Other dimensions of evaluation systems may also affect the acceptance or internalization of evaluation processes by performers. Future studies should seek additional information on the effects of frequency of evaluation and performer influence as well as explore the nature of other dimensions of evaluation systems.

The data from the six studies of the evaluation of teachers revealed a positive linear relationship between frequency of evaluation and teacher acceptance of the evaluation process and between teacher influence over the evaluation process and teacher acceptance of that process. No evidence of the negative effects of high levels of either variable was found to support the curvilinear relationships described in the propositions. This may be because there are no negative effects at any level of evaluation frequency or

performer influence or because the levels of evaluation frequency and performer influence in the schools in these six studies never approach the levels high enough to set in motion the predicted negative effect. Future studies might seek to identify schools where teacher evaluation is very frequent and those where teachers exercise high levels of influence over the evaluation process to explore the proposed curvilinear relationships more fully. Because administrator time is a scarce resource in many schools, it may be necessary to set up field experiments to obtain the conditions necessary to fully examine these predictions. This course of action would address a theoretical concern; from a practical standpoint, administrators might simply recognize that, in general, more frequent evaluation and greater teacher influence over the evaluation process will lead to greater teacher acceptance of the evaluation system.

Another aspect of the two propositions not addressed in the present analysis is the impact of differences in task predictability on the optimum level of evaluation frequency and performer influence to promote acceptance of the system. Future work should explicitly consider differences in the predictability of various teaching tasks and their impact on the proposed relationships.

Different approaches to data collection might permit investigators to more fully explore the implications of the two propositions. For example, the studies discussed here involved questionnaires in which teachers were asked to rate the relative frequency with which they experienced evaluation activities and the relative level of their influence over the evaluation process. Future studies might probe for the absolute levels of these variables as experienced by teachers. This might be accomplished (1) by questionnaires which present hypothetical situations to teachers and ask them to describe how their present situation compares along critical

dimensions to the hypothetical situations, (3) through interviews in which respondents are asked to more fully explain their answers to questions about the relative frequency of evaluations and their relative influence over the evaluation process, and (4) with observational studies which monitor the evaluation activities in schools.

Additional dimensions of the evaluation process should be explicitly considered in studies of the impact of evaluations on teachers. For example, teachers in the six studies reported here expressed concern about the reliability of the evaluations they received, that is, the extent to which the evaluations received by different teachers were comparable. They also expressed concern about the consistency of the evaluations conducted by different evaluators in a school system. This is of particular concern in those situations in which the results of evaluations are used by the school district to reduce the teaching force of the district.

One of the dimensions of evaluation systems treated as a dependent variable in the analysis above, the soundness of the evaluation system, may also function as an independent variable affecting the acceptability of the evaluation process to performers (Dornbusch and Scott, 1975). Thus, more frequent evaluations may lead to more soundly based evaluations (evaluations where the effort and performance level of the performer has more impact on the communicated evaluations), and more soundly based evaluations may lead to greater performer acceptance of the evaluation process.

Finally, the renewed interest in incentives for teachers suggests that the connections of evaluation systems to such incentives as well as the nature of the incentives themselves may be important dimensions of evaluation systems that have an impact on teacher acceptance of the evaluation process. In view of the relationship between such dimensions and the frequency of

evaluation and performer influence that appears when the results of the study in the district with the merit pay system are compared to results in other districts (Recall that teachers in the merit pay district reported more frequent evaluation activities and greater influence over the evaluation process), studies of the impact of incentive systems on teacher acceptance should carefully control these other dimensions of evaluation systems.

Evaluation processes are pervasive in all organizations. Given the enduring concern with improving the performance of teachers and students, developing an appreciation of the role of evaluation processes in schools should be high on the agenda of educational and social researchers. The theory of evaluation and authority has provided a conceptualization of the evaluation process that has guided a series of studies of evaluation systems as they affect both teachers and students. Further research based on this evolving theory should lead to the further development of the theory as well as to an enhanced understanding of the operation of educational organizations.

References

deCharms, R. and G. Natriello. "Evaluation and Teacher Motivation: Final Report to the Spencer Foundation." Washington University, St. Louis, Mo., 1981.

Dornbusch, S. M. and W. R. Scott. Evaluation and the Exercise of Authority. San Francisco: Jossey-Bass, 1975.

Dreeben, R. The Nature of Teaching. Glenview, Illinois: Scott, Foresman, 1970.

Duckworth, K. "Linking Educational Policy and Management with Student Achievement." Eugene, Oregon: Center for Educational Policy and Management, September 1981.

Landy, F. J., J. L. Barnes, and K. R. Murphy. "Correlates of Perceived Fairness and Accuracy of Performance Evaluation." Journal of Applied Psychology, 63, 1978:751-754.

Locke, E. A. "Personnel Attitudes and Motivation." In M. R. Rosensweig and L. W. Porter, (eds.) Annual Review of Psychology. Palo Alto, California: Annual Reviews, 1975.

McCauley, B.L. "Evaluation and Authority in Radical Alternative Schools and Public Schools." Unpublished Ph.D. Dissertation, Stanford University, Stanford, California, 1971.

McCauley, B.L., S.M. Dornbusch, and W.R. Scott. "Evaluation and Authority in Alternative Schools and Public Schools." Technical Report No. 23, Stanford Center for Research and Development in Teaching, Stanford, California, 1972.

McLaughlin, M.W. and D.D. Marsh. "Staff Development and School Change." Teachers College Record, 80, 1978:69-94.

Marram, G.D. "Visibility of Work and the Evaluation Process: Evaluation and Authority for Nurses in Hospitals and Teachers in Open and Closed Schools." Unpublished Ph.D. Dissertation, Stanford University, Stanford, California, 1972.

Marram, G.D., S.M. Dornbusch, and W.R. Scott. "The Impact of Teaming and the Visibility of Teaching on the Professionalism of Elementary School Teachers." Technical Report No. 33, Stanford Center for Research and Development in Teaching, Stanford, California, 1972.

Meyer, H. H., E. Kay, and J. R. French. "Split Roles in Performance Appraisal." Harvard Business Review, 43, 1965:123-129.

Natriello, G. and B. Rowe. "Life in a Loosely-Coupled World: The Evaluation of Elementary School Teachers." Unpublished Paper, Washington University, St. Louis, Mo., 1981.

Natriello, G. and S.M. Dornbusch. "Pitfalls in the Evaluation of Teachers by Principals." Administrator's Notebook. June 1981:1-4.

Natriello, G. and S.M. Dornbusch. Teacher Evaluative Standards and Student Effort. A Volume in the Institute for Research on Teaching Monograph Series. New York: Longman, in press.

Natriello, G. "Teachers' Perceptions of the Frequency of Evaluation and Assessments of their Effort and Effectiveness." American Educational Research Journal, forthcoming.

Natriello. G. and M. Cohn. "Beyond Sanctions: The Evolution of a Merit Pay System." Paper presented at the Annual Meeting of the American Educational Research Association, Montreal, Canada, 1983.

Szilagyi, A. D. "Causal Inferences Between Leader Reward Behavior and Subordinate Performance, Absenteeism, and Work Satisfaction. Journal of Occupational Psychology, 53, 1980:195-204.

Tannenbaum, A. S. Control in Organizations, New York: McGraw-Hill, 1968.

Thompson, J.E. "Evaluation and Authority in Elementary and Secondary Schools: A Comparison of Teachers and Administrators." Unpublished Ph.D. Dissertation, Stanford University, Stanford, California, 1971.

Thompson, J.E., S.M. Dornbusch, and W.R. Scott. "Failures of Communication in the Evaluation of Teachers by Principals." Technical Report No. 43, Stanford Center for Research and Development in Teaching, 1975.

Yerkes, R.M. and J.D. Dodson. "The Relation of Strength of Stimulus to Rapidity of Habit Formation." Journal of Comparative Neurology and Psychology, 18, 1980:459-482.